Clin Orthop Relat Res (2018) 476:612-614
DOI 10.1007/s11999.0000000000000157

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

CORR Insights

Published online: 13 February 2018

CORR Insights®: Validation of the Chinese (Mandarin) Version of the Oxford Knee Score in Patients with Knee Osteoarthritis

Thoralf Randolph Liebs MD, PhD

Where Are We Now?

There are several ways to measure treatment effects in orthopaedics. For example, arthroplasty registries traditionally consider the survival of the implant as the primary outcome [6]. In clinical trials, on the other hand, clinical scores such as the original Knee Society Score or the Hospital for Special Surgery scoring system are commonly used. Unfortunately, clinical scores might not always capture a patient's perception of his or her overall functional status, quality of life, and/or pain. In addition, many clinical scores require a physical examination, resulting in increased logistical burden and the possibility of observer bias. For these reasons, patient-reported outcome measures (PROMs), such as the Oxford Knee Score (OKS), have gained considerable attention [4].

Selecting an outcome measure is critically important when designing a clinical trial. For example, if a researcher chooses an outcome measure that is not responsive to clinical change, it is unlikely that such as study will identify treatment effects, even if the treatment delivers a "real" effect and the study design is otherwise perfect. An outcome measure should assess endpoints that matter most to patients, and it should be sufficiently defined so that it can be used by other researchers. In addition, robust outcomes tools should be well tested beforehand (unpublished and unvalidated outcome measures are associated with bias [8]), reliable (that is, it measures what it claims to measure), and reproducible so that there are no variations upon retesting [8].

These requirements might appear self-evident. However, there are examples of studies published in high-impact general medical journals where the disease-specific primary outcome was not validated [9] and even the algorithm for the calculation of the outcome was incomplete. It is not astonishing that such a study fails to detect differences between treatment groups [9]. Nevertheless, the study results influenced the recommendations of an important clinical practice guideline in orthopaedic surgery [11].

There are several PROMs to choose from when assessing knee osteoarthritis before and after arthroplasty. For example, both the WOMAC and the OKS are disease-specific instruments commonly used for the assessment of patients with osteoarthritis of the knee, and these instruments are considered reliable and reproducible. Although developed in English-speaking countries, several validated translated versions of these instruments are available in French, Spanish, and German.

Where Do We Need To Go?

Because of the increasing internationalization of clinical trials [12] and the large population in China, it would be advantageous if the results of TKA in China could be assessed by similar outcome measures that are used in Western countries. The authors of the current study strive to fill this gap. In their paper, Lin and colleagues translate and cross-culturally adapt the OKS to a Chinese (Mandarin) version and

This CORR Insights® *is a commentary on the article* "Validation of the Chinese (Mandarin) Version of the Oxford Knee Score in Patients with Knee Osteoarthritis" *by Lin and colleagues* DOI: 10.1007/s11999-017-5495-2.
The author certifies that neither he, nor any members of his immediate family, have any commercial associations (such as consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.
The opinions expressed are those of the writers, and do not reflect the opinion or policy of *CORR*® or The Association of Bone and Joint Surgeons®.

Thoralf R. Liebs MD (✉), PhD, Inselspital University Hospital, University of Bern, Freiburgstrasse 10, 3010 Bern, Switzerland, Email: Liebs@Liebs.eu

T. R. Liebs, Consultant Orthopaedic Surgeon, Inselspital University Hospital, University of Bern, Bern, Switzerland

calculate various psychometric properties for that score.

The Chinese version of the OKS could be considered a validated outcome instrument and can perhaps be used for assessing Mandarin-speaking patients undergoing TKA in China. However, as the current study validated the Chinese version of the OKS in terms of semantic equivalence and psychometric properties, it is not clear if it is appropriate to combine data collected by the Chinese version of the OKS with data collected by the original or other validated versions of OKS. Analyzing such measurement equivalence is important, since the assessment of differences among the outcome measures might be difficult to detect, thereby leading to misleading results and invalid findings [12].

The selection of the outcome measure influences the result of clinical comparative trials, and thereby influences treatment decisions. For example, it is possible, that a comparative randomized trial might demonstrate differences when outcome measure A is used, but no differences when outcome measure B is used. For this reason, it is mandatory that the primary outcome measure be selected in advance [8].

The Outcome Measures in Rheumatology Clinical Trials III Conference [2], which was ratified by the Osteoarthritis Research Society International Initiative (OARSI) Task Force [1] recommends using a core set of PROMS in osteoarthritis trials: Physical function, pain, and patient global assessment.

Currently, there is no clear recommendation as to which outcome measure specifically should be used in clinical trials of TKA. There are reasons, though, why joint-specific scales should be used, and that disease-specific tools alone are insufficient. For example, the OKS has been specifically developed to be "joint-specific in order to increase their sensitivity to the outcome of the joint replacement as far as possible and to be influenced as little as possible by other comorbidities" [10]. The OKS returns a single score result, in which both the pain dimension and the function dimension are included, as can be seen in the explanatory factor analysis presented by Lin as well. On the other hand, the WOMAC—a disease-specific tool for patients with osteoarthritis of the hip or knee, but not a joint-specific scale—was not designed initially to evaluate outcomes from total hip or total knee replacement, but it is commonly used for that purpose. The WOMAC returns four score results, one for each of the three subscores pain, function, and stiffness, and one for the overall score. Therefore, it might appear that the WOMAC is more appropriate to meet the recommendation of the OARSI. However, it has recently been demonstrated that the WOMAC "pain and function items that asked about the same activity consistently loaded on the same factor rather than on separate pain and function factors" [5]. Therefore, the strict differentiation between pain and function that is suggested when analyzing the pain and function subscores of the WOMAC has been questioned [5].

Determining which instrument to use depends on various factors, such as the setting of the research project. For example, national arthroplasty registries that aim to analyze PROMs in thousands of patients have different outcome selection criteria than initiators of randomized controlled trials.

How Do We Get There?

In order to compare the results of joint replacement surgery, standardization of the outcome measures is the key to success. The current study is important in this aspect, as it is the first step to implementing a validated Chinese version of the OKS that compares the success of knee replacement surgery in mainland China to other countries with different cultures.

Future studies, therefore, should assess whether there is measurement equivalence when comparing the Chinese version with the original version of the OKS. Measurement equivalence has been defined as present if "different language versions of a HRQoL instrument would yield similar scores at item and scale levels for respondents with identical levels of health-related quality of life" [7]. Several methods on how to assess measurement equivalence have also been discussed [12].

Although there is no gold standard on the horizon regarding the outcomes assessment of joint replacement surgery, it would be advantageous if clear recommendations were developed as to which outcome measure should be assessed in orthopaedic clinical trials, depending on the study design and other factors. These recommendations should be developed by expert panels in orthopaedics, specifically joint replacement surgery. We must reconsider the current trend of over-representing these panels with specialists who hardly ever see patients undergoing joint replacement surgery, such as epidemiologists, methodologists, internists, methods editors, patient representatives, health research methodologists, and biostatisticians [3, 11]. Once a recommendation has been developed, it should be published in scientific journals, mentioned in editorials, discussed among orthopaedic societies, and added to information for authors' guidelines. Furthermore, the administrators of www.clinicaltrials.gov could be asked to display a pop-up window showing information regarding the

recommended outcome measures as soon as a joint replacement randomized clinical trial is registered.

References

1. Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, Buckwalter J, Dougados M, Ehrlich G, Lequesne M, Lohmander S, Murphy WA, Jr., Rosario-Jansen T, Schwartz B, Trippel S. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Results from a workshop. *Osteoarthritis Cartilage*. 1996;4:217–243.
2. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, Altman R, Brandt K, Dougados M, Lequesne M. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol*. 1997;24:799–802.
3. Brignardello-Petersen R, Guyatt GH, Buchbinder R, Poolman RW, Schandelmaier S, Chang Y, Sadeghirad B, Evaniew N, Vandvik PO. Knee arthroscopy versus conservative management in patients with degenerative knee disease: A systematic review. *BMJ Open*. 2017;7: e016114.
4. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA*. 2013;309: 814–822.
5. Gandek B. Measurement properties of the Western Ontario and McMaster Universities Osteoarthritis Index: A systematic review. *Arthritis Care Res (Hoboken)*. 2015;67:216–229.
6. Liebs TR, Splietker F, Hassenpflug J. Is a revision a revision? An analysis of national arthroplasty registries' definitions of revision. *Clin Orthop Relat Res.* 2015; 473:3421–3230.
7. Luo N, Chew LH, Fong KY, Koh DR, Ng SC, Yoon KH, Vasoo S, Li SC, Thumboo J. Do English and Chinese EQ-5D versions demonstrate measurement equivalence? An exploratory study. *Health Qual Life Outcomes*. 2003;1:7.
8. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
9. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Ashton CM, Wray NP. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347: 81–88.
10. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, Dawson J. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br.* 2007;89: 1010–1014.
11. Siemieniuk RAC, Harris IA, Agoritsas T, Poolman RW, Brignardello-Petersen R, Van d V, Buchbinder R, Englund M, Lytvyn L, Quinlan C, Helsingen L, Knutsen G, Olsen NR, Macdonald H, Hailey L, Wilson HM, Lydiatt A, Kristiansen A. Arthroscopic surgery for degenerative knee arthritis and meniscal tears: A clinical practice guideline. *BMJ*. 2017;357:j1982.
12. Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, Hareendran A, Wiklund I, Chong LY, von Maltzahn R, Cohen L, Molsen E. Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: The ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. *Value Health*. 2009;12:430–440.